Age and Ageing 2003; **32:** 315–320

Identifying a cut-off point for normal mobility: a comparison of the timed 'up and go' test in community-dwelling and institutionalised elderly women

HEIKE A. BISCHOFF[1,2,3], HANNES B. STÄHELIN[2], ANDREAS U. MONSCH[2], MAURA D. IVERSEN[3,4], ANTJE WEYH[1], MARGOT VON DECHEND[5], REGULA AKOS[1], MARTIN CONZELMANN[6], WALTER DICK[1], ROBERT THEILER[5]

Departments of [1]Orthopedics and [2]Geriatrics, University of Basel, Switzerland
[3]Division of Rheumatology, Immunology and Allergy, the Robert B. Brigham Arthritis and Musculoskeletal Clinical Research Center, Brigham and Women's Hospital, Boston, MA, USA
[4]Doctoral Program in Physical Therapy, Graduate School for Health Studies, Simmons College, Boston, MA, USA
[5]Division of Rheumatology, Cantonal Hospital Aarau, Switzerland
[6]Department of Geriatrics, Felix Platter Spital, Basel, Switzerland

Address correspondence to: Heike A. Bischoff, Department of Medicine, Harvard Medical School, Division of Rheumatology, Immunology and Allergy, The Robert B. Brigham Arthritis and Musculoskeletal Clinical Research Center, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA. Fax: (+1) 617 731 9032.
Email: hbischof@hsph.harvard.edu

Abstract

Background: physical mobility testing is an essential component of the geriatric assessment. The timed up and go test measures basic mobility skills including a sequence of functional manoeuvres used in everyday life.

Objectives: to create a practical cut-off value to indicate normal versus below normal timed up and go test performance by comparing test performance of community-dwelling and institutionalised elderly women.

Setting and participants: 413 community-dwelling and 78 institutionalised mobile elderly women (age range 65–85 years) were enrolled in a cross-sectional study.

Measurements: timed up and go test duration, residential and mobility status, age, height, weight and body mass index were documented.

Results: 92% of community-dwelling elderly women performed the timed up and go test in less than 12 seconds and all community-dwelling women had times below 20 seconds. In contrast only 9% of institutionalised elderly women performed the timed up and go test in less than 12 seconds, 42% were below 20 seconds, 32% had results between 20 and 30 seconds and 26% were above 30 seconds. The 10^{th}–90^{th} percentiles for timed up and go test performance were 6.0–11.2 seconds for community-dwelling and 12.7–50.1 seconds for institutionalised elderly women. When stratifying participants according to mobility status, the timed up and go test duration increased significantly with decreasing mobility (Kruskall-Wallis-test: $p < 0.0001$). Linear regression modelling identified residential status ($p < 0.0001$) and physical mobility status ($p < 0.0001$) as significant predictors of timed up and go performance. This model predicted 54% of total variation of timed up and go test performance.

Conclusion: residential and mobility status were identified as the strongest predictors of timed up and go test performance. We recommend the timed up and go test as a screening tool to determine whether an in-depth mobility assessment and early intervention, such as prescription of a walking aid, home visit or physiotherapy, is necessary. Community-dwelling elderly women between 65 and 85 years of age should be able to perform the timed up and go test in 12 seconds or less.

Keywords: *mobility assessment, elderly, residential status, timed up and go test*

Introduction

The annual incidence of falls among community-dwelling persons increases from 25% at 70 years to 35% over the age of 75 [1, 2]. The consequences of falls are serious. Falls are the major source of injuries in the elderly, and are the 6th leading cause of death in individuals over 65 [3]. Falls are the most common causes of hip fractures and admission to nursing homes [4].

The accumulation of deficits in the elderly, such as muscular weakness, decreased balance or neuromuscular abnormalities, result in mobility impairments that may cause falls and difficulties in performing activities of daily living (ADL) [1, 5–8]. Functional mobility assessments enable physicians and other health care providers to set realistic, measurable goals for their patients, determine functional independence and assist in discharge planning. These tools mimic mobility efforts of everyday life, rather than assessing single neuromuscular impairments in elderly people [5].

The timed up and go test (TUG-test) is an effective method of assessing mobility and quantifying locomotor performance [9]. The TUG-test is objective, quick and easy to perform. The test includes basic mobility skills, such as rising from a chair, walking 3 meters, turning and sitting down on the same chair. The original description of the 'get-up and go' test by Mathias and colleagues [10] used a subjective rating scale. This test was later modified by Podsiadlo and Richardson, who developed a scaling system based on the observer's perception of the patient's fall risk [9]. The 'timed up and go version' made the scoring more objective, but the test remained unchanged. The original concept is appealing because it describes a realistic mobility assessment including potential fall situations, such as transfer in and out of a chair, walking and turning. Podsiadlo and Richardson validated the TUG-test in 60 elderly patients from a geriatric day hospital. They showed significant correlations (coefficient of correlation (r)=0.81) with the Berg Balance Scale [11], gait speed (r=−0.61) [12] and the Barthel Index (r=0.78) [13].

The purposes of this study were to create a practical cut-off value, below which TUG-test performance could be classified as normal and determine the influences of mobility status, age, gender, height, weight and body mass index (BMI) on TUG-test performance.

Methods

Subjects

Four hundred and ninety-one mobile elderly women were enrolled in this cross-sectional study, 413 community-dwelling (mean age in years 73.2 ± 3.2 SD) and 78 institutionalised elderly women (mean age 79.4 ± 3.7 SD). This ratio of community-dwelling and institutionalised elderly subjects was chosen to reflect the general population of elderly women in Switzerland.

Community-dwelling women were recruited from a centre in Aarau, Switzerland, which was currently running a prospective population-based cohort study [14]. Subjects screened were randomly selected from a population based community-dwelling cohort of postmenopausal women. Only 10% used a cane for longer walks, but all preferred to do the tests without a walking aid.

The institutionalised mobile elderly persons were recruited from two long-stay geriatric wards within acute geriatric care departments of the Felix Platter Spital and Kantonsspital in Basel, Switzerland. All institutionalised elderly women were awaiting nursing home placement and were medically stable.

The ethics committee of the Basel University Hospital approved the study protocol. Informed consent was obtained from subjects and legal guardians, where appropriate. An exclusion criterion was self-reported moderate or severe pain in weight-bearing joints at recruitment or examination. Persons with acute illness, hemiplegia or severe dementia, defined as being unable to consent or understand the instructions, were excluded from the study. Two major inclusion criteria were age (65–85) and ability to perform the TUG-test. Subjects who were either unable to walk 6 meters or unable get in or out of a chair with or without a walking aid were excluded.

Subjects were assigned to one of three physical activity groups adapted from a previous publication that related physical activity to bone remodelling in the elderly and showed a stepwise increase in bone resorption with increasing immobility [15].

Physical activity group 1 consisted of community-dwelling, independently living elderly. Community-dwelling elderly were able to walk safely outside and perform their own shopping. Physical activity group 2 and 3 included institutionalised elderly. Subjects assigned to activity group 2 were able to walk without a walking aid in the institution. Subjects assigned to activity group 3 used a cane or a walker. Institutionalised elderly rarely left the institution for a walk outside, and only if they were in company.

Raters

Three raters were used. The raters were a physician, one medical student and one physiotherapist. All raters were trained in the treatment of elderly persons. Each rater participated in a training session before conducting the study trials and received standardised instructions for the application of the test. A single person conducted all the tests on the same person.

Measures

This study used the TUG-test described by Podsiadlo and Richardson [9]. A special clinical examination room was used to perform all TUG tests. Subjects were observed and timed from the instant they rose from an armchair (seat height 48 cm; arm height 68 cm), walked 3 metres, navigated an obstacle on the floor (i.e. a brick placed a distance of 3 metres from the chair), and returned to a fully seated position in the chair. Subjects wore their regular footwear and were allowed to use the arms of the chair to get up.

Subjects began the test on the word, 'go' and were instructed to 'walk at a comfortable fast and secure pace'. Each subject performed the TUG-test three times on the same day, after one practice trial. The final score was the trial at which the subject was able to carry out the test the quickest of the three timed trials. The score is the time in seconds that the subject needed to complete the test. The test-retest reliability of the TUG-test performance scores was assessed on a random sample of 23 institutionalised patients. This subset of subjects was selected as the variability of TUG performance was expected to be greater than the variability of a subset of healthy women, increasing the likelihood of differences between testers. Spearman's rank correlation coefficients were used to assess inter-rater reliability (Inter-rater reliability=0.91; intra-observer reliability=0.96). All subjects who initially consented to join the study were willing to perform the TUG-test.

Medical comorbidity, which is likely to be associated with worse function, was measured using the Charlson Comorbidity index [16]. The medical charts of a subgroup of 50 institutionalised elderly women from one geriatric unit (Felix Platter Spital) were evaluated. The Charlson comorbidity index is a weighted index that takes into account the number and the seriousness of comorbid conditions. An index of '0' is associated with a 12% 1-year expected mortality rate, a score of '1–2' with a 26% 1-year expected mortality rate, a score of '3–4' with a 52% 1-year expected mortality rate and a score of '⩾5' with a 59% 1-year expected mortality rate [16].

Statistical analysis

Data were analysed by SPSS© computer program (Version 9; SPSS Inc., Chicago, IL). As data were slightly skewed, comparisons between residential groups were calculated with Mann–Whitney U tests. Difference in TUG-test performance between three mobility groups (community-dwelling, institutionalised with walking aid and institutionalised without walking aid) was calculated by Kruskal-Wallis-test. Alpha levels for significance was set at <0.05. With an alpha set at 0.05 and a beta set at 0.01, we calculated the estimated sample size needed using standardised response means. Based on these results, we more than exceeded the numbers in our sample to detect differences between groups on the TUG test performance scores.

The correlation coefficients were calculated by the Spearman rank sum method. To correct for age, gender and BMI, partial correlation coefficients were calculated.

In the multivariate analysis the following factors were included to explain total variation of TUG-test performance: residential status (community-dwelling *versus* institutionalised=mobility group 1 *versus* mobility group 2 and 3), physical activity score (community-dwelling plus institutionalised without walking aid *versus* institutionalised with walking aid), age, gender, height, weight and BMI. Comparing 10th and 90th percentiles of community-dwelling and institutionalised elderly established a normal cut-off point for TUG-test duration. Nonparametric 96.5% confidence intervals were calculated for the 10th and the 90th percentiles for community-dwelling and institutionalised elderly. Receiver operating characteristics (ROC) analysis was performed to evaluate the diagnostic validity of the TUG-test for discrimination of community-dwelling and institutionalised residential status.

Results

The primary objective of this study was to create a practical cut-off value, below which TUG-test performance could be classified as normal. The 10^{th}–90^{th} percentiles for TUG performance of community-dwelling elderly were 6.0 seconds [96.5% CI: 6; 6] to 11.2 seconds [96.5% CI: 11; 12] and 12.7 seconds [96.5% CI: 9; 14] to 50.1 seconds [96.5% CI: 38; 80] for institutionalised elderly. We therefore chose 12 seconds as a clinical cut-off point for normal TUG-test performance in community-dwelling elderly persons. With 96.5% confidence 12 seconds or less were regarded as normal. ROC analyses of all participants revealed a high diagnostic validity for discrimination of community-dwelling and institutionalised residential status by the TUG-test (Area under curve: 0.969) (see Figure 1). In a linear

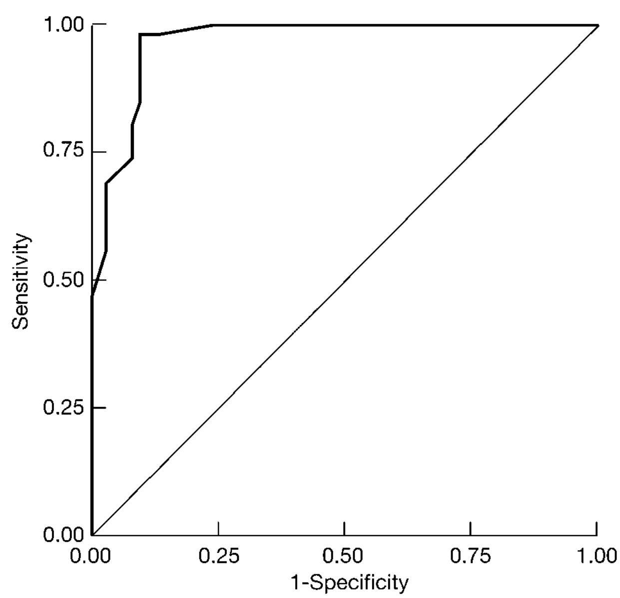

Figure 1. ROC – curve for TUG-test. Diagnostic validity of TUG-test performance in elderly women is shown giving the differentiation between community-dwelling and institutionalised elderly women. Area under the curve is 0.969. The positive predictive value is 0.76 and the negative predictive value is 0.98.

Table 1. Characteristics of study participants

Parameter (units)	Residential status (n)	Mean ($\pm$SD)	Median (interquartile range)	Minimum–maximum	p-value*
Age (years)	Community-dwelling (413)	73.2 ($\pm$3.2)	72 (71–75)	65–83	<0.0001
	Institutionalised (78)	79.7 ($\pm$3.7)	80.4 (69–82.7)	70–85	
Height (cm)	Community-dwelling (413)	159.2 ($\pm$5.6)	159 (155–163)	142–174	n.s.
	Institutionalised (78)	158.5 ($\pm$7.6)	159 (154.5–164)	140–175	
Weight (kg)	Community-dwelling (413)	67.7 ($\pm$12.4)	66 (58–75)	43–115	0.004
	Institutionalised (78)	62.3 ($\pm$12.9)	62 (53–72.5)	37.3–90	
BMI (kg/m^2)	Community-dwelling (413)	26.7 ($\pm$4.5)	26.1 (23.1–29.7)	17.9–42.4	0.004
	Institutionalised (78)	24.8 ($\pm$5.0)	24.3 (20.5–28.9)	16.2–37.5	
TUG (seconds)	Community-dwelling (413)	8.3 ($\pm$1.9)	8 (7–9.4)	4.8–15.8	<0.0001
	Institutionalised (78)	28.2 ($\pm$23.0)	21 (16–32.8)	8–160	

*Mann Whitney U-test.

regression model, including residential status, physical activity, age, gender, height, weight and BMI, only residential status ($p<0.0001$) and physical activity ($p<0.0001$) reached significance. The model explained 54% of total variation of TUG-test performance.

A second objective of this study was to compare TUG-test performance between community-dwelling and institutionalised elderly women. Table 1 displays the demographic characteristics and results of the TUG-test in community-dwelling and institutionalised elderly women. Institutionalised elderly were older ($p<0.0001$), lighter (weight: $p=0.004$) had a lower BMI ($p=0.004$) and were slower in the TUG–test ($p<0.0001$). The time scores on the TUG test ranged from 4.8–15.8 seconds in community-dwelling women and from 8–160 seconds in institutionalised women. In multivariate analysis, residential status and physical activity were the major predictors of TUG-test performance, after adjustment for age, height, weight and BMI. Figure 2 shows TUG-test performance in two age groups (65–74; 75–85 years). Regardless of residential status, age did not appear to influence TUG-test performance within groups (Mann–Whitney U test: n.s.).

A third objective of the study was to determine the influences of mobility status, age, gender, height, weight and BMI on TUG-test performance. Figure 3 shows TUG-test results in three mobility groups. Kruskal-Wallis-test on TUG-test performance was significantly different between mobility groups ($p<0.0001$). TUG-test results were significantly correlated with mobility (r=0.60; $p<0.0001$) and residential status (r=0.59; $p<0.0001$)

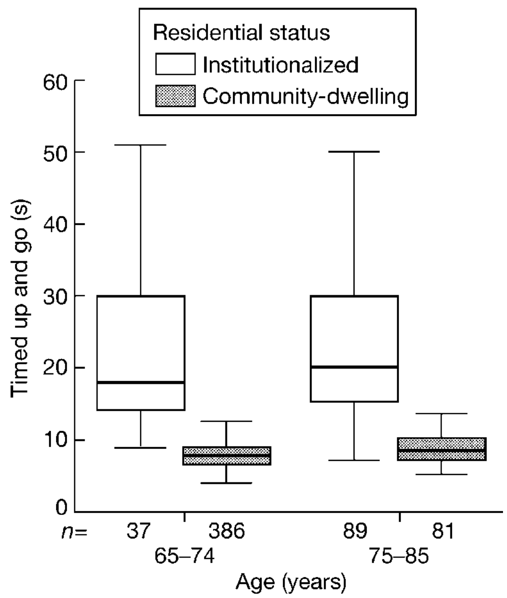

Figure 2. TUG-test performance of community-dwelling and institutionalised elderly women in two age groups. Within the two groups of institutionalised and community-dwelling elderly TUG-test performance was similar between age groups. Box plots show the median, inter-quartile range, and extreme cases of individual variables.

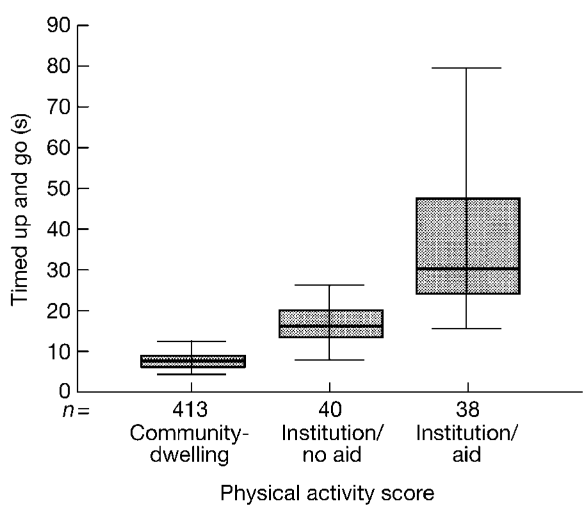

Figure 3. TUG-test performance in three mobility groups. TUG-test results are illustrated in three mobility groups. Kruskal-Wallis-test on TUG-test performance was significantly different between mobility groups ($p<0.0001$). Box plots show the median, interquartile range, and extreme cases of individual variables.

and both correlations stayed significant after correction for age and BMI (residential status: r=0.53; $p<0.0001$/mobility: r=0.65; $p<0.0001$).

Podsiadlo and Richardson [9] proposed three TUG-test performance groups depending on the duration of the test in seconds (group 1: <20 seconds; group 2: 20–30 seconds; group 3: >30 seconds). Figure 4 shows the distribution of community-dwelling and institutionalised elderly persons in four TUG-classes with an additional group that performed the TUG-test in <12 seconds. Ninety-two percent of community-dwelling elderly women performed the TUG-test in <12 seconds (74% <10 seconds) and all community-dwelling women had times below 20 seconds. In contrast only 9% of institutionalised elderly women performed the TUG-test in <12 seconds (8% <10 seconds), 42% were below 20 seconds, 32% had results between 20 and 30 seconds and 26% were above 30 seconds.

The univariate correlation between age and TUG-test duration was significant (r=0.45; $p<0.0001$) for the entire sample. However, if community-dwelling and institutionalised women were analysed separately, only community-dwelling elderly showed a significant correlation between age and TUG-test performance (r=0.24; $p<0.0001$). Among the community-dwelling women, a weak correlation was found between TUG-test duration and weight (r=0.18; $p<0.0001$) and BMI (r=0.21; $p<0.0001$). Height did not correlate significantly with TUG-test performance in either group.

Comorbid conditions, measured as a weighted index in a subgroup of 50 institutionalised elderly, did not correlate with TUG-test performance [16]. Two percent of the sample had a score of '0' indicating a 12% 1-year mortality rate, 56% had a score of '1–2' indicating a 26% 1-year mortality rate, 34% had a score of '3–4' indicating a 52% 1-year mortality rate and 8% had a score of '⩾5' corresponding to a 59% 1-year mortality rate.

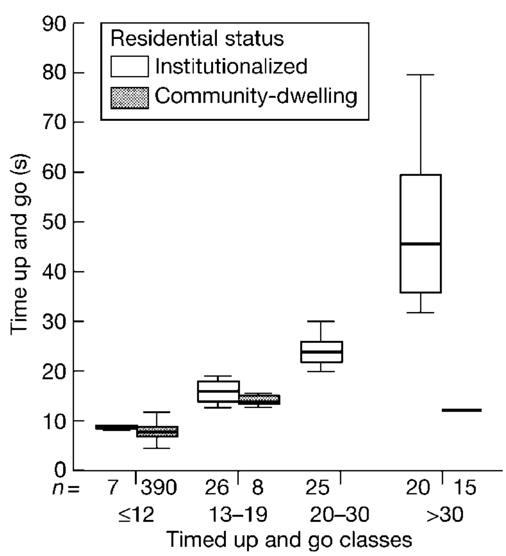

Figure 4. Distribution of study population in four TUG-test classes according to residential status. Distribution of community-dwelling and institutionalised elderly persons in four TUG-classes is shown. None of the community-dwelling women had TUG-scores above 20 seconds. Box plots show the median, interquartile range, and extreme cases of individual variables.

Discussion

In this study elderly women, ambulating safely outside, performed the TUG-test significantly quicker (10^{th}–90^{th} percentiles: 6.0–11.2 seconds), than mobile elderly women in the institution with or without walking aid (10^{th}–90^{th} percentiles: 12.7–50.1 seconds). Therefore, by comparing 10^{th}–90^{th} percentiles of both groups, we propose a cut-off point at 12 seconds or less as normal for TUG-test performance in this age group (65–85 years). This threshold showed discriminative value in the ROC-analysis with an area under curve of 0.969.

In Podsiadlo and Richardson's study [9] 10 healthy elderly persons were compared to 60 patients admitted to a geriatric day hospital. We were able to screen 491 community-dwelling elderly women in a population-based data set and compare them to 78 institutionalised elderly women. We confirm the findings of our colleagues, illustrating that most healthy community-dwelling elderly (77%) perform the TUG-test in <10 seconds.

We found that TUG-test duration increased in a stepwise fashion with decreasing mobility ($p<0.0001$). Both residential status and mobility status were identified as strong predictors of TUG-test performance by the adjusted analysis (r^2 0.54). Moreover correlation studies between TUG-test duration and residential and mobility status stayed significant after controlling for age and BMI. Age, as well as BMI, height and weight did not reach significance in the multivariate model. In bivariate correlation studies, age, BMI and weight showed weak correlation with TUG-test performance in community-dwelling elderly women, but not in institutionalised elderly women. Charlson comorbidity index [16], measured in a subgroup of institutionalised elderly women, did not correlate with TUG-test performance.

Our results suggest that TUG-test performance mirrors residential status in elderly women and discriminates between different mobility levels. One possible explanation for this result relates to the intent and design of the TUG-test which reflects physical performance during everyday life. It has been proposed that decline in physiological function may be better explained by the accumulation of deficits across multiple domains rather than by a single specific impairment (17–19). Accordingly, the TUG-test does not focus on independent effects of organ impairments, such as low muscle strength, decreased balance and other compairments, but measures the interaction of these factors on the performance of activities of daily living. It is supported by the growing emphasis on functional capacities in geriatric assessment.

In conclusion, we suggest that the TUG-test is useful in detecting mobility impairments in elderly persons. We believe the TUG-test is well suited to assess disability – it is easy to conduct, requires little equipment and has been shown to be reliable and valid. It clearly discriminates between community-dwelling and institutionalised elderly women. Furthermore, we propose a normative cut-off point at 12 seconds for elderly women between 65 and 85 years of age. In daily clinical practice, community-dwelling elders who perform the TUG-test in >12 seconds should receive early evaluation and intervention. Further studies are needed to validate this cut-off point against hard clinical outcomes such as falls and nursing home admissions.

Key points

- Twelve seconds is suggested as a practical cut-off value to indicate normal *versus* below normal timed up and go test performance.
- Community-dwelling elderly women should be able to perform the timed up and go test in 12 seconds or less.
- Timed up and go test performance is dependent on residential status.
- Timed up and go test performance decreases with impaired mobility.

References

1. Tinetti ME, Speechley M. Prevention of falls among the elderly. N Engl J Med 1994; 320: 1055–9.

2. Campbell AJ, Reinken J, Allan BC, Martinez GS. Falls in old age: a study of frequency and related clinical factors. Age Ageing 1981; 10: 264–70.

3. Baker SP, Harvey AH. Fall injuries in the elderly. Clin Geriatr Med 1985; 1: 501–12.

4. Tinetti ME, Williams CS. Falls, injuries due to falls, and the risk of admission to a nursing home. N Engl J Med 1997; 18: 1279–85.

5. Tinetti ME, Ginter SF. Identifying mobility dysfunction in elderly patients: standard neuromuscular examination or direct assessment. JAMA 1998; 259: 1190–3.

6. Tinetti ME, Inouye SK, Gill TM, Doucete JT. Shared risk factors for falls, incontinence, and functional depencence. JAMA 1995; 273: 1348–50.

7. Overstall PW, Exton-Smith AN, Imms FJ, Johnson AL. Falls in the elderly related to postural imbalance. Br Med J 1977; 1: 261–4.

8. Rantanen T, Guralnik JM, Sakari-Rantala R *et al.* Disability, physical activity, and muscle strength in older women: the women's health and aging study. Arch Phys Med Rehabil 1999; 80: 130–5.

9. Podsiadlo D, Richardson S. The timed "up & go": A test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991; 39: 142–8.

10. Mathias S, Nayak USL, Isaacs B. Balance in the elderly patient: the get-up and go test. Arch Phys Med Rehabil 1986; 67: 387.

11. Berg K, Wood-Dauphinee S, Williams JI *et al.* Measuring balance in the elderly: preliminary development of an instrument. Physiother Can 1989; 41: 304.

12. Wolfson SM, Whipple R, Amerman P, Tobin JN. Gait assessment in the elderly: A gait abnormality rating scale and its relation to falls. J Gerontol 1990; 45: 12–19.

13. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. MD State Med J 1965; 14: 61.

14. Hans DB, Thiebaud D, Lippuner K *et al.* Can we compare the ultrasound reference data for the calcaneum with the DXA data for the hip and spine?: a Swiss population based study. J Bone Miner Res 1999; 14: 257.

15. Theiler R, Stähelin HB, Kränzlin M, Tyndall A, Bischoff HA. High bone turnover in elderly people. Arch Phys Med Reha 1999; 80: 485–9.

16. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987; 40: 373–83.

17. Duncan PW, Chandler J, Studensky S, Hughes M, Precot B. How do physiological components of balance affect mobility in elderly men? Arch Phys Med Rehabil 1993; 74: 1343–9.

18. Boonen S, Aerssens J, Breemans S, Dequeker J. Fractures of the proximal femur: Implications of age-related decline in muscle function. J Orthop Rheumatol 1995; 8: 127–33.

19. Coupland C, Wood D, Cooper C. Physical inactivity is an independent risk factor for hip fracture in the elderly. J Epidemiol Commun Hlth 1993; 47: 441–3.

Received 10 April 2001; accepted in revised form 18 December 2002